



A HOZZÁADOTTÉRTÉK-MODELLEK ALKALMAZÁSA A TANULÓI TELJESÍTMÉNYMÉRÉSEKBEN

Balázs Ildikó

Oktatási Hivatal, Köznevelési Programok Főosztálya

Az utóbbi tíz évben a tanulóiteljesítmény-mérések felhasználási területe jelentős mértékben bővült. Az oktatási rendszerek sorra ismerik fel és használják ki az oktatási rendszer egészének kiadásaihoz viszonyítva csekély anyagi ráfordítást igénylő és objektív adatokkal szolgáló méréseket abból a célból, hogy mind az oktatásirányítást, mind a közoktatási rendszer szereplőit (fenntartók, iskolák és pedagógusok), mind annak felhasználóit (szülők és tanulók) informálják a rendszer egészének vagy egy-egy szegmensének, szereplőjének bizonyos szempontok szerinti teljesítményéről (IEG, 2006; UNESCO, 2007; Clarke, 2011).

A McKinsey & Company két nagy hatású oktatási tárgyú jelentése kimutatja, hogy egyrészt a fejlett és hatékony oktatási rendszerek hatékony külső mérési-értékelési rendszerrel rendelkeznek, melyeknek a tanulók eredményeinek mérése az egyik pillére (Barber & Mourshed, 2007). Másrészt az elmúlt évtizedben az eredményeiken jelentős mértékben javító oktatási rendszerek közös jellemzője, hogy a tanulók teljesítménymérési rendszerének fejlesztése szerepelt a reformok között, ami arra utal, hogy a rendszer aktuális eredményességétől függetlenül az eredményesség növeléséhez a mérési-értékelési rendszer kiépítése és fejlesztése elengedhetetlen (Mourshed, Chijioko, & Barber, 2010).

A külső mérési-értékelési rendszerek egyik fontos kérdése, hogyan értelmezi és milyen összefüggésekben vizsgálja a tanulói eredményeket. A mérési rendszer fejlesztésének egyik lehetséges iránya éppen a tanulói teljesítmények értelmezésének árnyaltabb megközelítése. Az első mérési-értékelési rendszerek a tanulók teszteredményeit, nyers pontszámait vették figyelembe, a tanulók korábbi eredményét, szociális, kulturális és gazdasági helyzetét figyelmen kívül hagyva. Azonban a nyers pontszám önmagában nem a legjobb indikátora a tanárok, iskolák eredményességének: először Erik Hanushek (1971) tanulmányában jelent meg az a megközelítés, ami a tanárok munkájának hatékonyságát úgy méri, hogy a tanulók korábbi eredményeit, az eredményre gyakorolt családi hatást és az osztálytársak hatását is figyelembe veszi az aktuális eredményük értékelésekor. Az ilyen típusú hozzáadottérték-modellek (*added value models*) sokkal alkalmasabbak arra, hogy egy-egy iskola vagy tanár teljesítményéről reális képet adjanak. Éppen ezért az elmúlt egy-két évtized szakmapolitikai paradigmaváltása során – a tényeken alapuló szakmapolitika igényének előtérbe kerülésével – e modellek is kiemelt figyelmet kaptak (Braun, Chudowsky, & Koenig, 2010; OECD, 2008). Az elmúlt egy-két évtized szakirodalma részletesen fog-

lalkozik a modellek technikai megvalósításának lehetőségeivel, a statisztikai eljárások elméleti és gyakorlati hátterével (pl. Lissitz, 2005, 2006) éppúgy, mint a hozzáadottérték-modellek bevezetésének oktatáspolitikai lépéseivel és következményeivel (pl. OECD, 2008).

Jelen tanulmány a hozzáadottérték-modellek szakirodalmába nyújt betekintést, röviden ismerteti az e modellekhez kapcsolódó szakmai dilemmákat, valamint bemutatja a magyar mérési gyakorlatban megjelenő modelleket. A tanulmány első részében röviden vázoljuk az tanulóieredmény-modellek típusait, és definiáljuk, mit értünk hozzáadottérték-modellel alatt. A második rész ismerteti néhány, az elmúlt évtizedben e témában született nemzetközi tanulmánykötetet és szakmai konferenciát. A harmadik rész a hozzáadottérték-modellek bevezetése során felmerülő szakmai kérdéseket és megoldandó feladatokat sorolja fel, míg a negyedik részben a lehetséges statisztikai modelleket mutatjuk be. Az ötödik rész a magyarországi előzményeket és az Országos kompetenciamérésben jelenleg alkalmazott gyakorlatot ismerteti. A tanulmány a hozzáadott érték (*added value*) iskolai szintű felhasználásának lehetőségeit tekinti át, nem tér ki részletesen a tanári szintű hozzáadottérték-modellek jellemzőire és alkalmazhatósági feltételeinek ismertetésére.

A tanulóieredmény-modellek típusai

A gyakorlatban az oktatási intézmények vagy tanárok értékelésére használt teszt alapú értékelési modelleket a mérési eredmények kezelése és az eredmények interpretálása szempontjából négy nagyobb csoportba sorolhatjuk: státuszmodellek, a kohorszok összehasonlításán alapuló modellek (*cohort-to-cohort change models*), növekedési modellek és hozzáadottérték-modellek (Braun, Chudowsky, & Koenig, 2010).

A *státuszmodellek* a tanulók teljesítményének egy időpontban mért értékét állítják az elemzés középpontjába, gyakran összehasonlítva azt egy előre meghatározott célértékkel. A státuszmodellek egy speciális változata az, ami az eredmények értékelésénél a tanulók különböző, az iskola által nem befolyásolható szocio-kulturális háttéradatait figyelembe veszi. Ezeket a modelleket nevezzük *kontextuális eredményességi modelleknek*. A *kohorszok összehasonlításán alapuló* modellek az időről időre elvégzett mérések eredményeinek összehasonlításán alapulnak, például két egymást követő tanévben mért eredményeket vetik össze, azt vizsgálják, változott-e a tanulók eredménye. Ebből következtetnek arra, hogy javult-e az oktatás minősége az eltelt időben.

A *növekedési modellek* a tanulók eredményeit követik, és azt vizsgálják, mennyit fejlődtek a tanulók a két mérés között eltelt időszakban. Ezek a modellek már csak longitudinális mérési rendszerek esetében alkalmazhatók, ahol a tanulók két vagy több mérési időpontban mért eredményei összeköthetők.¹ Ugyanakkor a növekedési modellek nem vizsgálják, hogy a teljesítményben bekövetkezett változás milyen mértékben tulajdonítható a különböző hatásoknak, milyen mértékben függ a teljesítménynövekedés az iskolán

¹ Néhány mérési rendszerben akkor is alkalmazzák a növekedési modelleket, ha csak az iskolát nem váltó tanulók eredményeit tudják összekötni. Ha a két mérési időpont között az iskolát váltó tanulók száma kellően alacsony, a rendszernek ez a hiányossága csak kismértékben torzítja az eredményeket.

kívüli és az iskola befolyása alá tartozó tényezőktől, csupán a növekedés mértékére koncentrálnak. A *hozzáadottérték-modellek* ezzel szemben megpróbálják az iskola hatását elkülöníteni az egyéb, az iskola által nem befolyásolható tényezőktől. Szintén a tanulók két mérési időpont közötti fejlődésének mértékét vizsgálják, ám nem a fejlődés mértéke áll a figyelmük középpontjában, hanem arra a kérdésre keresik a választ, hogy egy-egy iskola tanulói teljesítményének fejlődésében mekkora szerepet játszott az adott iskola, mekkora a tanulók teljesítményének növekedése ahhoz a fejlődési potenciálhoz képest, amit egy „átlagos, tipikus” iskolában sikerült volna elérniük ugyanezeknek a tanulóknak a vizsgált időszakban.

A hozzáadottérték-modellek definíciója

A szakirodalomban számos definíciót olvashatunk a hozzáadott értékre és a hozzáadott-érték-modellekre vonatkozóan, melyek közös pontja, hogy a hozzáadottérték-modellek az iskola, illetve a tanár vagy egy oktatási program, módszer, intézkedés hatását próbálják becsülni a tanulói teljesítményekre. E modellek a tanuló eredményességét meghatározó bonyolult összefüggésrendszeren belül a vizsgált tényező hatását igyekeznek elkülöníteni az attól független egyéb tényezők hatásától (Value-added Measures, 2013). A hozzáadott érték modellek a tanulók tudásának fejlődéséből kiindulva, esetleg további, az iskola által nem befolyásolható tényezők hatását is figyelembe véve becsülik az iskola, tanár munkájának eredményességét.² A hozzáadott érték e tanulmányban használt általános definíciója a következő: [„Az iskola által hozzáadott érték] az iskola hozzájárulása a tanulók fejlődéséhez az oktatás valamely meghatározott, előírt céljának (pl. a kognitív képességek fejlődésének) teljesülése érdekében. Az iskola hozzájárulása alatt az egyéb, a tanulók fejlődésére ható tényezők figyelembe vétele melletti nettó hatás értendő. ...[A hozzáadottérték-modellek] a statisztikai modellek egy olyan osztálya, amely az iskola hozzájárulását becsüli a tanulók fejlődéséhez az oktatás valamely meghatározott, előírt céljának (pl. a kognitív képességek fejlődésének) teljesülése érdekében legalább két időpontban végzett mérés alapján (OECD, 2008. p. 17.).

A hozzáadottérték-modellek az oktatási rendszerek azon jellemzőjét igyekeznek a statisztika módszereivel kiküszöbölni, hogy a tanulók iskolákba tagozódása nem randomizált, véletlen beosztással történik, ahol minden iskola várhatóan ugyanolyan eséllyel indulna a gyermekek fejlesztése terén, és minden, a tanulók eredményeiben mutatkozó különbség az iskola hatásának volna tulajdonítható (ahogyan ez egy különböző kezelési módszereket összehasonlító tervezett kísérlet során lenne). Az iskolarendszer ezen jellemzője miatt az iskolák eredményességének összehasonlítására a tanulók eredményeinek

² A tanulmányok egy része a hozzáadott érték becsléséhez a tanuló egy pontban mért eredményét használja fel, kontrollálva a tanuló családi háttérének hatását a teszteredményére, az iskola hozzáadott értékét a regresszió reziduálisával becsülve (Csapó, 2002; Balázs & Zemléni, 2004). E modelleket jelen tanulmányban a kontextuális eredményességi modellek körébe soroljuk, és csak azokat a modelleket tekintjük hozzáadottérték-modelleknek, amelyek a tanuló legalább két pontban mért eredményét felhasználják, és az adott idő alatti fejlődés mértékéből becsülik az iskola, tanár vagy program hatását

egyszerű összevetése nem megfelelő, hiszen az például régóta közismert és többször bizonyított tény, hogy a tanulók szocio-kulturális háttere igen erősen korrelál eredményeikkel, és legtöbbször még a teljesítményük növekedésével is, így az iskola eredménye is nagymértékben függ attól, milyen a tanulói szocio-kulturális összetétele (McCall, Kingsbury, & Olson, 2004).

A kontextuális eredményességi modellek ennek a hatásnak a kiküszöbölésére születtek, ugyanakkor az iskola eredményességének megítélésére mégsem jelentenek általános megoldást, hiszen a tanulóknak ezekben a modellekben vizsgált egyetlen teszteredménye nemcsak a jelenlegi iskolájuktól és tanáraiktól, valamint a családi háttérüktől függ, hanem a korábbi nevelési-tanulási tapasztalataiknak összegzése az óvodától kezdődően (OECD, 2008; Harris & Sass, 2005). A hozzáadottérték-modellek ezzel szemben az iskolán kívüli hatásoknak és a tanuló eredményét befolyásoló korábbi hatásoknak az együttes kiküszöbölésére tesznek kísérletet (Stringfield & Yakimowski-Srebnick, 2005).

Hozzáadottérték-modellek a nemzetközi gyakorlatban

A hozzáadottérték-modellek célja az iskolák (pl. tanárok, programok) eredményességének, minőségének jellemzése, és ebből eredően a hozzáadottérték-modellek első megjelenése az iskola hatékonyságának, a tanárok hatásának vizsgálata kapcsán merült fel (OECD, 2008). A nemzetközi szakirodalom, főleg az Egyesült Államok állami mérési-értékelési programjaihoz kapcsolódóan, nagyon sok kísérleti vagy már a gyakorlatban működő hozzáadott érték típusú értékelési rendszerről számol be. A legnagyobb és legrégebbi ezek közül az Oktatási Hozzáadott Érték Elszámoltathatósági Rendszer (*Educational Value-Added Accountability System*, EVAAS), amit először Tennessee államban alkalmaztak az 1990-es évek közepén, akkor még *Tennessee Value-Added Accountability System* (TVAAS) néven. Jelenleg a SAS Institute által kifejlesztett hozzáadottérték-módszert Tennessee államon kívül Ohio és Pennsylvania államokban, valamint az Egyesült Államok sok más iskolakörzetében is használják (Sanders, Saxton, & Horn, 1997; Ballou, Sanders, & Wright, 2004; Wright, Sanders, & Rivers, 2006; Betebenner & Linn, 2010). A teljesítménymérések szakirodalmában a TVAAS-rendszer fejlesztői használták először, honosították meg és terjesztették el szélesebb körben a hozzáadott érték fogalmát.

A hozzáadottérték-modellek főbb jellemzőinek korai összegzését adja Carol Taylor Fitz-Gibbon 1997-es jelentése. A brit kormány felkérésére azt vizsgálta, milyen lehetőségeket rejtenek, milyen előnyökkel járnak és milyen feltételekkel és jellemzőkkel vezethetők be a hozzáadottérték-modellek az iskolák eredményességének vizsgálatára az Egyesült Királyságban. A jelentés konklúziója szerint a hozzáadottérték-modell bevezetése az iskolák eredményességének mérésében statisztikailag korrekt, intuitív, könnyen érthető és elfogadatható mérőszámot eredményez (Fitz-Gibbon, 1997; Bognár, 2000). Angliában 2002 óta tartalmaznak hozzáadottérték-eredményeket az iskolajelentések az iskolaszaka-

szok végén letett vizsgák eredményeiből kiindulva (Ray, 2006). A táblázatokban felhasznált adatokat és az alkalmazott hozzáadottérték-modelleket időről időre felülvizsgálják és amennyiben szükséges, módosítják.³

A hozzáadottérték-modellek körül nemzetközi szinten jelenleg is számos kutatás és fejlesztés zajlik, a 2000-es években több olyan kutatási konferencia volt, ahol e modellek módszertanát vizsgálták. Ezeken a konferenciákon az ilyen modellek terén végzett legfrissebb kutatásokat mutatták be, a módszerek alkalmazhatóságát és korlátait vitatták meg a kutatók, és az előadásokat, eredményeket konferenciakötetekben vagy a konferencia web-lapján meg is jelentették. Ilyen volt például az a két Marylandi Egyetemen tartott konferencia, amelyek előadásainak anyagát Robert W. Lissitz szerkesztette kötetbe. Az első konferencia 2004. október 22- és 23-án zajlott *Conference on Value Added Modeling* címmel (Lissitz, 2005); a másodikra 2005. november 7-én és 8-án került sor *Longitudinal Modeling of Student Achievement Conference* címmel (Lissitz, 2006). A két konferencia előadásai a hozzáadottérték-modellek előtt álló legfontosabb módszertani kihívásokat írják le, azokra keresnek választ, új módszereket, modelleket ismertetnek.

Hasonló kérdésekkel foglalkoztak a Wisconsin-Madison Egyetem 2008. április 22–24. között megtartott konferenciáján (*National Conference on Value-Added Modelling*)⁴, illetve a 2009 decemberében az *Educational Testing Service (ETS) Center for K-12 Assessment and Performance Management* részlege által szervezett szemináriumon. Az *Exploratory Seminar: Measurement Challenges Within the Race to the Top Agenda* egyik fő témája a tanulói teljesítmények növekedésének szummatív mérési módszerei volt (ETS, 2010). Mindhárom konferencia résztvevői egyetértettek abban, hogy a hozzáadottérték-modellek hasznos eszközök lehetnek az iskolák és a tanárok munkájának megítélésében, azonban az alkalmazott tesztek, modellek és elemzések megválasztása kellő körültekintést igényelnek, és nagymértékben befolyásolják a kapott eredmények megbízhatóságát, relevanciáját és felhasználhatóságát.

A módszertani konferenciák mellett olyan konferenciák, műhelybeszélgetések is voltak az elmúlt néhány évben, amelyek a hozzáadottérték-modellek oktatáspolitikai vonatkozásait, az erre alapuló elszámoltathatósági rendszerek lehetőségeit és korlátait, a bevezetésükhöz kapcsolódó várt és váratlan, kívánatos és nemkívánatos hatásokat vitatták meg. Ezen konferenciák eredményeként többnyire olyan kiadványok jöttek létre, amelyek a módszertani eszköztár áttekintése mellett az ilyen elszámoltathatósági rendszerek létrehozásának feltételeit, előnyeit és veszélyeit is tárgyalják. Ezek közül az egyik legnagyobb hatású az *OECD Added Value* kutatási projektje volt, melyben az OECD-országok e területen dolgozó vezető szakemberei vitatták meg a hozzáadott érték típusú elszámoltathatósági rendszerekben rejlő lehetőségeket, és amelynek eredményeként az OECD ajánlásokat és módszertani útmutatókat is megfogalmazó kötetet adott ki (OECD, 2008). Hasonló szempontok vezették az egyesült államokbeli Nemzeti Kutatási Tanácsot (*National Research Council*) és Nemzeti Oktatási Akadémiát (*National Academy of Education*),

³ Az egyes évek iskolai teljesítménytáblázataiban szereplő adatokról lásd a <http://www.education.gov.uk/schools/performance/archive/index.shtml> oldalt.

⁴ A konferencia programja és prezentációi elérhetők az egyetem honlapján http://www.wcer.wisc.edu/news/events/natConf_papers.php

amikor megszervezték a hozzáadottérték-modellekkel foglalkozó műhelykonferenciát, melynek jelentése szintén a hozzáadottérték-modellek körüli legfontosabb módszertani kérdéseket járja körül, de az OECD-jelentéssel szemben kifejezetten az Egyesült Államok jelenlegi törvényi szabályozásának és mérési-értékelési rendszereinek szempontjából vizsgálja a kérdést (Braun, Chudowsky, & Koenig, 2010). A következő fejezet az e kötetben tárgyalt legfontosabb oktatáspolitikai és módszertani kérdéseket foglalja össze.

A hozzáadottérték-modellek módszertani kérdései

A megalapozott hozzáadottérték-modelleknek számos kihívással és módszertani nehézséggel kell megküzdeniük, melyeket hat nagyobb kérdéskörbe csoportosíthatunk (Meyer & Dokumaci, 2010; Braun, Chudowsky, & Koenig, 2010; OECD, 2008; Lissitz, Doran, Schafer, & Willhoft, 2006): (1) a hozzáadottérték-modelleket alkalmazó tanulóiteltjesítmény-mérési rendszerek céljai és a hozzájuk kapcsolódó következmények; (2) a mérőeszközök és a belőlük számított képességpontok jellemzői; (3) a tanulói eredmények követése, az évek közötti összekötés minősége; (4) az adatfelvételek minősége és az adatok mennyisége; (5) az alkalmazott hozzáadottérték-modellek jellemzői; (6) az eredmények interpretálása, bemutatása.

A rendszer céljai és a hozzá kapcsolódó következmények

A hozzáadott érték alapú mérési rendszer által elérhető célok tisztázása, a rendszer előnyeinek és korlátainak hangsúlyozása, a következmények nagyságának, a mérés tétjének meghatározása (*high-stakes* vs. *low-stakes*), a rendszerhez kapcsolódó következmények és ösztönzők kiépítése elengedhetetlenül az első kérdések között kell szerepeljen a hozzáadottérték-rendszer tervezésekor (Braun, Chudowsky, & Koenig, 2010; OECD, 2008). Az OECD 2008-as elemzése három fő alkalmazási területet különböztet meg (OECD, 2008; *Chapter one – Policy Objectives of the Development of a System of Value-Added Modelling*): *az iskola eredményességének növelését szolgáló kezdeményezések adatokkal való ellátása, az iskola elszámoltathatóságának biztosítása, valamint az iskolaválasztás informálása*. Ez a három alapvető cél a hozzáadottérték-modell fejlesztése során részben eltérő döntéseket eredményezhet. Az iskola belső teljesítményértékelésének és a minőség növekedésének biztosítása céljából létrehozott rendszer például nem igényli az adatok nyilvánosságra hozását, ám az adatok értelmezéséhez szükséges iskolai kapacitás kiépítését feltételezi. Ugyanakkor ahhoz, hogy a rendszerből származó adatok megfelelően informálják a szülők és tanulók iskolaválasztását, nem csak az adatok nyilvánossága szükséges, hanem az is, hogy ezek az adatok, elemzések könnyen elérhetőek és könnyen értelmezhetőek legyenek, valamint a médiában is megfelelő figyelmet kapjanak (OECD, 2008).

A legtöbb hozzáadottérték-modellt alkalmazó mérési rendszer elszámoltathatósági céllal jött létre, a jelen tanulmányban példaként bemutatott amerikai mérések mind az iskolák

elszámoltathatóságának eszközei. Ezek a rendszerek nem kizárólag hozzáadottérték-modelleket használnak – az elszámoltathatósági rendszer fogalmáról, céljairól, típusairól és az alkalmazott modellekről l. Tóth (2010) összefoglalóját.

Az angliai hozzáadottérték-mutatók fő felhasználási területei négy nagyobb csoportba sorolhatók (Ray, 2006): (1) teljesítmény-táblázatokba rendezve megbízható és összehasonlítható adatok biztosítása minden iskola eredményességéről a szülők és a szélesebb közvélemény számára, az iskolák elszámoltathatóságának biztosítása; (2) adatok szolgáltatása az iskolai fejlesztő munkához és a szakmai ellenőrzések számára; (3) iskolák kiválasztása bizonyos oktatási programokban, kezdeményezésekben való részvételre; (4) oktatáspolitikai kezdeményezések eredményességének monitorozása. Itt tehát az említett célok mindegyike megjelenik valamilyen formában, a nyilvánosság és a szakmai ellenőrzés informálása éppúgy, mint az eredményesség növelését szolgáló kezdeményezések kiszolgálása.

A mérőeszközök és a belőlük számított képességpontok jellemzői

A hozzáadottérték-modellek alapjaként alkalmazott mérőeszközök relevanciája, megbízhatósága és érvényessége, a teszteredmények rezisztanciája a csalással és a tesztpontszám inflációjával szemben, a pontszámok mérési hibája minden teszt alapú értékelési rendszernél fontos kérdés, de a hozzáadottérték-modelleknél, főleg, ha az az egyes tanárok eredményességéről akar információkat adni – ahogyan az az Egyesült Államokban sok esetben történik –, különösen nagy fontossággal bírnak (Braun, Chudowsky, & Koenig, 2010). A mérési hiba mértéke a növekedési modelleknél és a hozzáadottérték-modelleknél kiemelt szerepet játszik, hiszen a két vagy több mérési pontban mért értékek hibája kumulálódik. A hozzáadottérték-modellek esetében további torzítást okozhat az is, hogy a korábbi mérési eredmény(ek)e)t a regressziós modell független változójaként, változóiként alkalmazzuk, amit általában hibamentesnek feltételezünk (Zumbo & Forer, 2011; Doran & Jiang, 2006; McCafrey, Lockwood, Koretz, & Hamilton, 2003; Ladd & Walsh, 2002).

A különböző mérési időpontokban kapott eredmények összekötése

Az egyes időpontokban, életkorokban felmért képességpontok statisztikai jellemzői és összekötése, a pontszámok összehasonlíthatósága, az alkalmazott képességskála vertikálitása, a tesztekben meglévő esetleges plafonhatás a hozzáadottérték-moddell bevezetése során szintén vizsgálandó kérdés. A legtöbb ilyen modell olyan vertikális teljesítményskálák (*vertical scaling*) használatát igényli, amely a különböző időpontokban mért tanulói eredményeket ugyanazon az intervallumskálán helyezi el (Doran, Jiang, Cohen, Gushta, & Phillips, 2005; Kolen & Brennan, 2004; Ballou, Sanders, & Wright, 2004). Ez korántsem triviális feltétel, az intervallumskálák készítésével kapcsolatos nehézségek (Ballou, 2009; Briggs & Betebenner, 2009) mellett a közös skála kialakítása sem magától értetődő, a lehetséges tesztfelépítési döntések és az alkalmazott ekvivalenciát biztosító módszerek pedig nagymértékben befolyásolhatják a hozzáadottérték-modellek eredményeit és az eredmények megbízhatóságát (Briggs, Weeks, & Wiley, 2008; Patz, 2007).

Az adatfelvétel minősége és az adatok mennyisége

A modell számításához felhasznált adatok mennyisége és minősége, a tanulók követésének megbízhatósága, az adatvesztésből eredő torzító hatások szintén befolyásolják az eredmények megbízhatóságát. Azokban az esetekben, amikor a hozzáadottérték-modelleket a tanárok értékelésében akarják felhasználni, a tanuló-tanár összekötési adatok megléte és megbízhatósága is kérdéses. Az adatok mennyisége és minősége minden teszteredményeket felhasználó értékelési rendszerben fontos kérdés, ugyanakkor a növekedési modelleknél és a hozzáadottérték-modelleknél az adatok összekötése új elem, amely tovább növelheti az adathiányt. Az iskolánkénti, osztályonkénti alacsony tanulószámmal járó problémája így ebben az esetben fokozottan jelentkezik (Braun, Chudowsky, & Koenig, 2010; Braun, 2005).

Az egy-egy iskolában felmért tanulók száma befolyásolja az iskola hozzáadott értékére kapott becslés megbízhatóságát és stabilitását. Az évről évre nagymértékben ingadozó hozzáadott értékek, amelyek mögött nem az iskola működésében bekövetkezett tényleges változás, csak statisztikai bizonytalanság áll, alááshatja a rendszer elfogadottságát. Ezért az ilyen ingadozások vizsgálata különösen fontos, és a modell kidolgozása során döntést kell hoznunk arról, hogy a tanulók száma vagy a becslés megbízhatósági tartományának nagysága alapján mely iskolákat zárunk ki az elemzésből. Az OECD szakértői csoportja – saját korábbi tapasztalatai alapján – azoknak az iskoláknak az esetében, ahol 20–30-nál kevesebb tanuló adata használható a hozzáadott érték számításához, a becslés hibája túlságosan nagy lehet. Ugyanakkor az iskolák mérete és a modellek hibája minden oktatási rendszerben más és más lehet, ezért minden új hozzáadottérték-modell bevezetésénél meg kell vizsgálni az adatok mennyiségének kérdését (OECD, 2008).

Az alkalmazott hozzáadottérték-modell jellemzői

A hozzáadott érték számítása során alkalmazott modell jellemzői, az egyes modellek alkalmazásával kapott értékek relevanciája és megbízhatósága, a becslésekhez kapcsolódó hibák nagysága, a becslések precizitása és stabilitása, a lehetséges torzító hatások szintén fontos kérdések. A hozzáadottérték-modellek az általános vegyes lineáris regressziós modellek (*general mixed linear models*) változatai, ahol a korábbi pontszám, a tanulói és az iskolai kontextus mint változók különböző módon, fix vagy véletlen hatásként kerülnek be a modellbe (OECD, 2008; Lissitz et al., 2006).

A gyakorlatban számos különböző konkrét hozzáadottérték-modellt alkalmaznak, melyek a modell alapfeltevéseiben és a bevont változók körében is különbözhetnek. Az OECD 2008-as tanulmányának nyomán, szemléltetésként, a hozzáadottérték-modellek három főbb csoportját vázoljuk: a *lineáris regressziós modellek*, a varianciakomponens-modellek vagy *hierarchikus lineáris modellek*, valamint a *fixhatás-modellek* általános leírását adjuk meg. E modellek mellett léteznek a *többváltozós véletlenhatás-modellek*, valamint a *növekedésgörbe-elemzések* is, ugyanakkor ezek a variánsok iskolánként kettőnél több mérési eredményt igényelnek, így a jelenlegi hazai mérési rendszerben nem igazán alkalmazhatók.

Az egyszerű lineáris regressziós modell

Az egyszerű lineáris regressziós modellekben a tanuló pontszámát lineáris regressziós egyenlettel becsüljük a korábbi mérésben elért pontszáma és néhány háttérváltozója alapján. A j . iskola i . tanulója a következő egyenletet írhatjuk fel:

$$y_{ij(2)} = a_0 + a_1 \cdot y_{ij(1)} + \sum_{k=1}^P b_k \cdot X_{kij} + \varepsilon_{ij} \quad (1)$$

ahol $y_{ij(1)}$ és $y_{ij(2)}$ a két egymást követő mérési eredményt jelöli, X_{kij} a modellben figyelembe vett tanulói szintű kontextuális változók, a_k , b_k jelöli a regressziós együtthatókat és végül ε_{ij} a független, normális eloszlású, állandó szórásnégyzetű reziduális (OECD, 2008).

Ennél a modellnél az egy-egy tanuló esetében a hozzáadottérték-bebecslésünk a második teszten elért tényleges és a modell alapján becsült pontszám különbsége lesz; egy-egy tanulócsoportra, osztályra, iskolára a hozzáadottérték-bebecslés a csoporthoz tartozó tanulók hozzáadottérték-bebecsléseinek átlaga. A modell könnyen alakítható az adott mérés tulajdonságaihoz, a figyelembe vett korábbi eredmények köre és a kontextuális változók köre rugalmasan alakítható (Ladd & Walsh, 2002; Jakubowsky, 2008). E modell előnye, hogy könnyen értelmezhető, és a tanulói reziduálisok vagy akár a modell együtthatóinak közlése elegendő ahhoz, hogy az iskolában is egyszerűen kiszámíthassák egy-egy tanulócsoport esetében a hozzáadott értéket. A hozzáadottérték-modell 2006-os lengyelországi bevezetésénél kiterjedt képzési programot alkalmaztak annak érdekében, hogy a modell lényegét a tanárok megértsék, illetve maguk is tudjanak bebecsléseket végezni saját tanulóik esetében (OECD, 2008).

A hierarchikus lineáris modellek

A varianciakomponens-modellek (vagy hierarchikus lineáris modellek) két regressziós egyenlet kombinációi, melyek segítségével a modellben kapott reziduális értékek varianciáját két szintre, az iskolák közötti és az iskolán belüli tanulók közötti részre bontjuk szét. Ennek a modellnek az előnye, hogy a tanulók iskolákba rendezett voltát, az adatbázis „hierarchikus” jellegét is figyelembe veszi. A modell legegyszerűbb változata (OECD, 2008):

$$y_{ij(2)} = a_{0j} + a_1 \cdot y_{ij(1)} + \sum_{k=1}^P b_k \cdot X_{kij} + \varepsilon_{ij} \quad (2)$$

$$a_{0j} = a_0 + \delta_{0j} \quad (3)$$

ahol $\varepsilon_{ij} \sim N(0, \sigma^2)$ és $\delta_{0j} \sim N(0, \tau^2)$ független normális eloszlású változók. Ebben a modellben az iskolai hozzáadott érték bebecslései nem csupán az iskola saját tanulóinak adataitól függenek, az egy-egy iskolára az iskola hozzáadott értékének legkisebb négyzetek módszerével (*ordinary least square*) számított bebecslése a bebecslés hibájának nagyságától függően egy-egy korrekciós tényezőt kap, amivel a bebecslés „eltolódik” az iskolai szinten

átlagos hozzáadott értéket reprezentáló a_0 felé. Ezt a jelenséget hívják zsugorító (*shrinkage*) hatásnak, hiszen az iskola hozzáadott értékét kisebbíti. Mivel az iskolai hozzáadott érték becslésének hibája függ a tanulók számától, a kisebb iskolák esetében a zsugorító hatás általában nagyobb lesz (OECD, 2008; Lissitz et al., 2006). Az iskolai szintű reziduálisok normális eloszlással való modellezése miatt szokás a hierarchikus lineáris modelleket véletlenhatás-modelleknek nevezni.

A modell bővíthető egyrészt a figyelembe vett szintek száma szerint, az iskola-tanuló felosztás további bontásával könnyen általánosítható az iskola-osztály-tanuló háromszintes felosztásra. Másrészt beépíthetünk a modellbe iskolai szintű változókat is, melyek a (3) egyenletben jelennek meg független változókként; valamint a tanulói szintű független változók, a korábbi teszteredmény és a tanulói háttérváltozók is szétválaszthatók az iskolai átlagok és az iskolán belüli tanulói szintű hatások együttesére (Bryk & Raudenbush, 1992) – a hierarchikus lineáris modellek kompetenciamérésben való alkalmazásának egy példáját l. Balázs és Zempléni (2004) tanulmányában.

A fixhatás-modellek

A fixhatás-modellek a véletlenhatás-modellekkel szemben az iskolai hozzáadott értéket úgy építik be az egyenletünkbe, hogy minden egyes iskolára egy-egy plusz változót definiálunk, ami az adott iskolába tartozó tanulók esetében 1, minden más tanulóra 0. Ennek a változónak a modellből kapott együtthatója lesz az adott iskola hozzáadott értéke. Az (1) egyenlet módosításával ez a modell a következőképpen írható fel:

$$y_{ij(2)} = a_0 + a_1 \cdot y_{ij(1)} + \sum_{k=1}^P b_k \cdot X_{kij} + \theta_j + \varepsilon_{ij} \quad (4)$$

ahol θ_j a j . iskola hozzáadott értéke.

A fixhatás-modellben – a véletlenhatás-moddellel szemben – nincs korrekciós tényező, az iskola hozzáadott értékének becslése csak az iskola tanulóinak adataitól függ, így ennél a modellnél a hozzáadott értékek torzítatlan becsléseit kapjuk. Ugyanakkor a becslés hibájának nagysága sem csökken, így az ezzel a modellel kapott becslések évről évre való ingadozása a véletlenhatás-modellekhez képest nagyobb lehet (OECD, 2008; Lissitz et al., 2006).

A fixhatás- és véletlenhatás-modellek kombinációját alkalmazzák az egyik legismertebb egyesült államokbeli hozzáadott érték alapú rendszerben, amit Dallas (Texas állam) iskoláiban alkalmaznak. A kétlépcsős eljárásban először egyszerű lineáris regressziós modellel megbecsülik a tanulók pontszámait minden mérésre a kontextuális változók alapján, majd az ebből a modellből kapott standardizált reziduálisokra kétszintes hierarchikus lineáris modellt alkalmaznak az iskolák hozzáadott értékének becsléséhez. Ez utóbbi modellben a tanulói szintű egyenletben a korábbi teszteredmények előző modellből kapott standard reziduálisai, az iskolai szintű modellben pedig különböző iskolai szintű háttérváltozók, például az iskola szocio-kulturális összetétele szerepelhetnek független változókként (Webster, 2005; Webster & Mendro, 1997).

A többváltozós véletlenhatás-modellek és a növekedésgörbe-elemzések

A többváltozós véletlenhatás-modellek alkalmazásához több mérési területen és lehetőleg minél több alkalommal mért teszteredményekre van szükség, ahol a tanulók által elért eredményeket többváltozós válasz-moddellel (*multivariate response model*) írhatjuk fel, ahol a tárgyaként és egyes időpontokban kapott reziduálisok normális eloszlásúak pozitív definit kovariancia mátrixszal (OECD, 2008; Lissitz et al., 2006). A modellek jellemzője, hogy az iskolai hatások el nem évülő hatásként jelennek meg, a korábbi évek hatásai ebben a modellben minden későbbi teszteredményben változatlan súllyal szerepelnek.

A többváltozós véletlenhatás-modellek legismertebb változata az EVAAS (Wright, White, Sanders, & Rivers, 2010). Az EVAAS-modellek többnyire nem vesznek figyelembe a modelljeikben kontextuális változókat, az iskola hatása csak a tanulók pontszámából származik, ugyanakkor erre egyrészt a modell kidolgozói szerint nincs is szükség, mert a tanulónkénti több mérési területről és több évből származó adat miatt minden tanuló a saját kontrolljaként jelenik meg a modellben (erről bővebben l. Sanders, Wright, Rivers, & Leandro, 2009). Másrészt a modell EVAAS-C nevű variánsa alkalmaz kontextuális változókat is, az ott alkalmazott eljárás mintaként szolgálhat hasonló modellek kidolgozására (Sanders, Saxton, & Horn, 1997; Ballou, Sanders, & Wright, 2004). Előnye az ilyen típusú modelleknek, hogy egy-egy hiányzó teszteredmény esetén sem törlik a tanuló teljes adatsorát, a meglévő teszteredményeket ekkor is fel tudják használni a számítások során. Azonban a modell hazai gyakorlatban való alkalmazhatóságát korlátozza a számításokhoz szükséges több időpontban felvett többféle teszteredmény, ami a kompetenciamérés esetében nem áll rendelkezésünkre, az összesen iskolánként két mérési pont ehhez nem elegendő (a 6. és a 8. évfolyam az általános iskolához rendelhető, a 8. és a 10. évfolyam a középiskola kezdeti és elért eredményének tekinthető). Ugyanakkor a HELP program esetében a modell alkalmazása elképzelhető, ott ugyanis sokféle kognitív képességet mérnek évenkénti adatfelvétellel, ami bőséges alapot biztosíthat a modellek futtatásához (a kompetenciamérésről és a HELP programról később lesz szó). Hasonlóan, legalább három mérési eredmény szükséges a növekedésgörbe-elemzések alkalmazásához, ami feltételezi, hogy a tanulók teszteredményeinek időbeni változása egy mögöttes képességnövekedési görbének köszönhető, és ezt alapul véve megbecsüli, hogy az iskola hozzájárulása nyomán ez a görbe hogyan változott meg, mennyiben gyorsította vagy lassította az iskola hatása a tanuló képességeinek fejlődését (OECD, 2008; Ponisciak & Bryk, 2005).

A különböző modellek közötti választás nem magától értetődő, noha a szakirodalom számos olyan összehasonlításról számol be, ahol a különböző modellek alkalmazása nagyon hasonló, magasan korreláló eredményeket hozott (Ponisciak & Bryk, 2005; Webster, 2005); ugyanakkor a komplexebb modellek, valószínűleg, robosztusabbak a hiányzásokból, az adatvesztésből és a mérési hibákból adódó torzításokkal szemben. Fontos, hogy a könnyű érthetőség és a bonyolult modellek alkalmazása közötti egyensúlyt megtaláljuk.

A modell típusának kiválasztásánál is fontosabb, hogy milyen tanulói és iskolaszintű változókat veszünk figyelembe a hozzáadott érték számításakor, ugyanis ez a döntés nagymértékben befolyásolhatja a modellben kapott iskolai reziduális értékek, azaz a becsült

hozzáadott értékek mértékét (Tekwe et al., 2004). A bevont változók köre okozhatja például az iskola vagy tanár hatásának szisztematikus alul- vagy felülbecslését, hiszen ha a tanulók családi háttere nem kerül be a modellbe független változóként, akkor könnyen előfordulhat, hogy az iskolák tanulói összetételéből eredő fejlődésbeli különbségek hozzájárulnak az iskola becsült hozzáadott értékéhez. Ezzel szemben, ha a családi háttér hatását figyelembe vesszük, akkor előfordulhat, hogy a kapott hozzáadott érték alulbecsüli az iskola hatását, amennyiben az iskola valódi hozzájárulása és a tanulók családi háttere között valamiféle kapcsolat áll fenn. Ez utóbbi eléggé természetes feltételezés, hiszen a legtöbb iskolarendszerben a jó tanárokért folyó versenyben a jobb tanulói összetételű iskolák előnyben vannak, a tanári kar összetétele pedig befolyásolja az iskola tényleges hozzáadott értékét.

Az eredmények interpretálása, bemutatása

Az eredmények bemutatásának, interpretálásának és felhasználásának a módjáról szintén a tervezés során kell döntenünk. Ennek során különös figyelmet kell fordítani arra, hogy a felhasználás és az interpretálás módja illeszkedjék a meghatározott célokhoz, a rendelkezésre álló adatokhoz és az alkalmazott módszerekhez. Fontos, hogy az eredményeket úgy mutassuk be, hogy az a célközönség számára könnyen értelmezhető legyen, ugyanakkor pontos, a félreértés esélyét minimálisra csökkentő szemléltető ábrák, táblázatok és leírások szülessenek, melyek egyaránt felhívják a felhasználók figyelmét az eredmények értelmezésének lehetőségeire és korlátaira (OECD, 2008).

Az eredmények ismertetésekor fontos szempont, hogy milyen közönséget kívánunk megszólítani. Más megközelítésre, más adatokra és leírásokra lehet szükség attól függően, hogy az eredményeket az oktatásirányítás, az iskolák és tanárok, a szülők és tanulók vagy a szélesebb közvélemény számára szeretnénk bemutatni. Az eredmények interpretálásának fontos részei az adatokat bemutató írott anyagok mellett a célközönség(ek) számára szervezett, az eredmények értelmezését és felhasználását segítő rendszeres előadások és bemutatók, személyes vagy online képzések.

Hozzáadottérték-modellek a magyar mérési gyakorlatban

A magyar mérési gyakorlatban a 2000-es évek elején honosodott meg a pedagógiai hozzáadott érték fogalma. A kifejezést az e tanulmányban elfogadott jelentése mellett igen gyakran alkalmazták, alkalmazzák az itt kontextuális eredményességi modelleknek nevezett modellekre is, melyek a tanuló és az iskola eredményességét nem a korábbi eredmény, hanem kizárólag az iskola által nem befolyásolható szociális, gazdasági és kulturális jellemzők figyelembe vételével ítélik meg. Ilyen értelemben használja a kifejezést Csapó Benő (2002) tanulmányában is, ami a hozzáadott érték mérésének egyik első magyar kísérletét ismerteti; a vizsgálatba bevont osztályok eredményeit vetették össze a szülők iskolai végzettségének figyelembe vételével. Balázs és Zempléni (2004) szintén a családi

háttér alapján várt és tényleges teljesítmény összevetésével vizsgálta az iskolák hozzáadott értékét a 2003. évi Országos kompetenciamérés adatainak felhasználásával.

Mivel Magyarországon a családi háttér tanulói teljesítményre gyakorolt hatása nemzetközi összehasonlításban is igen magas (Balázs, Ostorics, Szalay, Szepesi, & Vadász, 2013), az ilyen típusú elemzések fontos szemponttal járulnak hozzá az iskolák eredményességének megítéléséhez: az iskola hátránykompenzáló hatását mutatják be. Ugyanakkor ezek a kontextuális eredményességi modellek nem képesek a felvételi szelekció hatását kiszűrni, ugyanolyan eredményesnek mutathatnak két iskolát akkor is, ha az egyik a jó eredményt tényleges pedagógiai munkával érte el, a másik viszont a felvételi eljárása során az ugyanolyan családi háttérű gyerekek közül a tehetségesebbeket, az eleve jobb képességeket válogatta ki.

Éppen ezért fontos, hogy az iskola hozzáadott értékét az e tanulmányban használt értelemben, a tanulók korábbi eredményének és jelenlegi eredményének kapcsolatát összevetve is vizsgáljuk. Kevés olyan adatbázis, kutatás van a magyar oktatási rendszerben, amely a tanulók eredményeit egyéni szinten is követhető módon, több időpontban is vizsgálja, így alkalmas ilyen típusú elemzések végzésére. A három legnagyobb, jelenleg is zajló longitudinális, iskolás gyerekeket, az iskolai tanulási folyamatokat vizsgáló program a *Szegedi Iskolai Longitudinális Program (HELP)*, az *Életpálya-felmérés*, illetve az *Országos kompetenciamérés (OKM)*.

A Szegedi Iskolai Longitudinális Program

A HELP az MTA-SZTE Képességfejlődés Kutatócsoport programja, ami az első átfogó longitudinális iskolai felmérés Magyarországon (Csapó, 2007, 2014). A program 2003-ban indult három, a 9., az 5. és az 1. évfolyamos tanulók közül országosan reprezentatív módon választott mintával, majd négyévenként újabb 1. évfolyamos mintával egészül ki a kutatás, tehát a kezdeti három populációhoz 2007-ben, majd 2011-ben és 2015-ben újabb minta kapcsolódott. A vizsgálat céljai között szerepel, többek között, az iskola kezdő szakaszának vizsgálata, az iskolai kudarcok, lemaradások okainak feltérképezése, a tanulók kognitív képességeinek fejlődésére ható tényezők feltérképezése, a tanulók egyéni eredményei és az őket befogadó osztályok, iskolák teljesítményei közötti kapcsolatok feltárása. A HELP első adatainak elemzése is leginkább ezekre a területekre fókuszálnak, például Tóth, Csapó és Székely (2010) megállapítja, hogy az iskolák és az osztályok közötti teljesítménykülönbségek tendenciaszerűen nőnek a magasabb évfolyamok felé haladva. A HELP adatai alkalmasak a hozzáadott érték típusú elemzések végzésére is, ugyanakkor a program elsődleges célja nem ez, és ilyen elemzések eddig nem készültek.

Az Életpálya-felmérés

Az Életpálya-felmérés a TÁRKI és az Educatio közös projektje. Kertesi Gábor kutatásvezető irányításával tízezer fiatal életét követik nyomon 2006 ősztől évi rendszerességgel (Kertesi & Kézdi, 2008, 2010, 2012; Hajdú, Kertesi, & Kézdi, 2014). A kutatás fő célja az iskolai sikertelenség és lemorzsolódás okainak feltérképezése, a tanulmányi sike-

resség és a későbbi társadalmi beilleszkedés és boldogulás szempontjából különösen veszélyeztetett csoportok életút-vizsgálata. A felmérés alapja a 2006 májusában 8. osztályos tanulók sokasága, akik kitöltötték az Országos kompetenciamérés szövegértés és matematika tesztjeit, valamint a családháttér-kérdőívet. A kiinduló minta 10 000 fős volt, a kompetenciamérésben alacsony eredményt elért és a sajátos nevelési igényű tanulók a kutatás céljai miatt felülreprezentáltak, ami azonban az egyes tanulók életútjának követése érdekében stratifikált egylépcsős mintavételi eljárással készült (Kézdi, 2009), ezért nem alkalmas az iskolai hozzáadott érték típusú elemzések elvégzésére.

Az Országos kompetenciamérés

Az Országos kompetenciamérés 2008-tól követi egyéni szinten a tanulókat a mérési azonosító bevezetésével, és 2010-től ad hozzáadott érték típusú elemzéseket minden iskola számára. Magyarországon ez az egyetlen mérés, ami minden iskolára vonatkozóan éves szinten szolgáltat ilyen adatokat. A kompetenciamérésben minden év májusában a matematikai eszköztudást és a szövegértési képességet vizsgáló tesztek, valamint tanulói, telephelyi és intézményi kérdőívek kitöltésével vesznek részt a 6., a 8. és a 10. évfolyamos tanulók és iskoláik. A teszten elért eredmény alapján a tanulók tesztelméleti modellel számított képességpontokat kapnak, a képességskála mindhárom évfolyamra és a 2008 óta eltelt évekre is közös, összehasonlítható. A tanulói, telephelyi és intézményi szintű eredményekről jelentés készül, a tanulói eredményeket a tanuló, szülei és az iskolája ismerhetik meg, a telephelyi és az intézményi szintű eredmények nyilvánosak, az internetről letölthetők.⁵ A mérés jellemzőiről és eredményeiről az Oktatási Hivatal Közoktatási Mérési Értékelési Osztályának kiadványai (Balázsi et al., 2014; Balázsi, Lak, Szabó, Szabó, & Vadász, 2014; OH-KMÉO, in press) teljes körű tájékoztatást adnak.

A 2010. évi telephelyi jelentés már tartalmazott hozzáadottértékmutatókat, és ezeket azóta is minden évben közzéteszi az Oktatási Hivatal (OH-KMÉO, 2014). A kompetenciamérés célja elsősorban az iskolák tájékoztatása tanulóinak eredményességéről a két felmért területen oly módon, hogy eredményeiket az ország összes iskolájához és a hozzájuk valamilyen szempontból hasonló iskolákhoz viszonyíthassák. Éppen ezért a jelentés tervezésekor a cél nem egyetlen eredményességi mutató kialakítása volt, hanem a tanuló és az iskola eredményének a lehető legtöbb szempont szerinti elemzése. Ezért a jelentésekbe 2008-ban a már korábban is feltüntetett átlageredmény és annak a tanulók családi hátterével való összevetése mellett három különböző hozzáadott érték típusú elemzés került.

Egyszerű lineáris regresszió

A legegyszerűbb mutató a telephelyen tanulók adott évi mérésben elért átlageredményének a két évvel korábbi átlageredményükkel való összevetése a telephelyi átlagokra illesztett lineáris regresszió alapján (OH-KMÉO, 2014, in press). Ezzel az egyszerű lineáris regressziós módszerrel minden telephelyre megadható, hogy a tanulók átlageredményének két év alatti változása esetükben gyengébb, hasonló vagy erősebb, mint amennyit

⁵ <https://www.kir.hu/okmfit/>

az ország egy átlagos iskolájában változott volna az átlageredmény, feltéve, hogy a telephelyen kellően sok tanulóra van adatunk mindkét érintett mérésből. A mérési azonosító segítségével a telephely mérés kori tanulóinak a két évvel korábbi eredménye akkor is hozzárendelhető az adott telephelyhez, ha időközben a tanulók iskolát váltottak; ahogyan a 10. évfolyamos tanulók többségére ez igaz is.

A korábbi eredményen alapuló hierarchikus lineáris modell

A jelentésben bemutatott másik két hozzáadott érték típusú elemzés a korábban leírt hierarchikus lineáris modelleken alapul. Ezek közül az első csak a tanulók korábbi eredményét és az adott mérésben elért eredményét veszi figyelembe, sem tanulói, sem iskolai szintű háttérváltozókat nem tartalmaz (OH-KMÉO, 2014, in press). Ugyanakkor a modell figyelembe veszi, hogy a korábbi eredmény telephelyen belüli hatása, illetve a tanulók átlageredményének telephelyek közötti hatása más lehet, ezért a korábbi eredményt két szintre bontva építi a modellbe. Emellett nem fix együtthatókat alkalmaz, hanem az együtthatók is véletlen változók, azaz a modell több, összesen négy részre bontja a reziduális értéket, és a következő egyenletekkel adható meg:

$$y_{ij(2)} = a_{0j} + a_{1j} \cdot (y_{ij(1)} - y_{j(1)}) + a_{2j} \cdot (y_{j(1)} - y_{..(1)}) + \varepsilon_{ij} \quad (5)$$

$$a_{0j} = a_0 + \delta_{0j} \quad (6)$$

$$a_{1j} = a_1 + \delta_{1j} \quad (7)$$

$$a_{2j} = a_2 + \delta_{2j} \quad (8)$$

ahol $y_{ij(1)}$ és $y_{ij(2)}$ a j . telephely i . tanulójának két évvel korábbi és az adott mérésben elért eredménye, $y_{j(1)}$ a j . telephely „korábbi átlageredménye”⁶, $\varepsilon_{ij} \sim N(0, \sigma^2)$ és

$\begin{pmatrix} \delta_{0j} \\ \delta_{1j} \\ \delta_{2j} \end{pmatrix} \sim N(0, \Omega_\delta^2)$ független normális eloszlású véletlen változók a reziduális értékek, az

utóbbi 3x3-as kovariancia mátrixszal. A tanulói szintű ε_{ij} hibatag mellett telephelyi szinten így még három, egymással akár korreláló hibatagot enged meg a modell. Ez egyrészt azt jelenti, hogy a tanulók korábbi eredményének telephelyen belüli hatása eltérhet az egyes telephelyek esetében, a δ_{1j} hibatag miatt a telephelyen belüli becslések regressziós egyenesei nem feltétlenül párhuzamosak. Másrészt a δ_{1j} és δ_{2j} hibatag miatt a telephelyi szintű reziduális értékek szórása a tanulók korábbi eredménye és a telephelyek „korábbi

⁶ Itt a (2) időpontban a j . telephelyen tanuló diákok két évvel korábbi átlageredménye szerepel, függetlenül attól, hogy a tanulók az (1) időpontban ugyanezen vagy más telephelyen tanultak-e.

átlageredménye” függvényében változhat (vagyis a modell az egyszerű lineáris modellekkel szemben nem követeli meg, hogy a reziduális értékek szórása a független változók minden lehetséges értékére ugyanannyi legyen).

A 2013. évi Országos kompetenciamérés 10. évfolyamos és a 2011. évi mérés 8. évfolyamos matematika tesztjeinek eredményeire illesztett modell együtthatói a következők voltak (az együtthatók után zárójelben a hibák szerepelnek):

$$a_0 = 1665(1,3), \quad a_1 = 0,55(0,005), \quad a_2 = 0,99(0,010), \quad \sigma^2 = 12811(157,8),$$

$$\Omega_{\delta}^{-2} = \begin{pmatrix} 2223(150,9) & & \\ 4,44(0,379) & 0,03(0,002) & \\ -6,23(0,709) & -0,003(0,003) & 0,04(0,007) \end{pmatrix}.$$

Az együtthatók értékéből látható, hogy a telephely tanulóinak korábbi átlageredménye nagyobb hatással volt a tanulók 10. évfolyamos eredményére, mint a tanuló saját eredményének a telephelyi átlagtól való eltérése ($a_2 > a_1$). A Ω_{δ}^{-2} kovariancia mátrix egyrészt azt mutatja, hogy az együtthatók mellé rendelt hibatagok mindegyike szignifikáns, azaz javítják a modell illeszkedését a tényleges adatokra. Másrészt δ_{0j} és δ_{1j} között a korreláció pozitív, δ_{0j} és δ_{2j} között pedig negatív (δ_{1j} és δ_{2j} között nincs szignifikáns korreláció). Az előbbi következtében a magasabb hozzáadott értékkel rendelkező telephelyek esetében (ahol δ_{0j} nagyobb) általában többet számít a tanuló saját korábbi eredménye, a telephelyen belüli regressziós egyenes meredekebb. Az utóbbi azt mutatja, hogy a magasabb korábbi átlageredményű telephelyek esetében a reziduálisok szórása kisebb.

A jelentésben e modell alapján olyan ábra szerepel, amely a tanulók korábbi és jelenlegi eredményét mutatja be az adott telephelyre, feltüntetve az e pontokra illesztett egyszerű regressziós egyenest, valamint a modell alapján becsült regressziós egyenest is. Itt tehát nem egyetlen, a telephely hozzáadott értékét jellemző számot, hanem egy ennél jóval több információt tartalmazó ábrát kapnak a telephelyek. Hasonló, csak háromszintes – telephely, osztály és tanuló – modell alapján az egyes osztályok tanulóinak fejlődéséről is képet kaphatunk a korábbi eredményt is három szintre bontva, ugyanakkor a modell egyszerűbb értelmezhetősége és stabilitása érdekében a háromszintes modellben csak fix együtthatók szerepelnek (OH-KMÉO, 2014, in press).

Komplex hozzáadottérték-mutató

A harmadik hozzáadott érték típusú mutató, amit a jelentések tartalmaznak, már tanulói, osztály- és telephelyi szintű háttérváltozókat is figyelembe vesz (OH-KMÉO, 2014, in press). A modellben a tanulók, valamint jelenlegi osztályuk és telephelyük korábbi eredménye mellett a tanuló neme, családjának vagyoni helyzete (a tanulói háttérkérdőív alapján számított index), az otthonukban található könyvek száma, a szülők legmagasabb iskolai végzettsége (egyéni, osztály- és telephelyi szinten is), valamint a telephely településtípusa és képzési formája szerepelnek magyarázó változókként. A modell meglehetősen

komplex, igyekszik a tanuló fejlődésére ható minél több iskolán kívüli tényezőt figyelembe venni. Ezzel a tanuló reziduális értéke már az iskolán kívüli hatásoktól jórészt megtisztított érték lesz, és mint ilyen, valószínű, hogy jobban becsüli az iskola munkájának eredményességét.

A modell bonyolultsága miatt a telephelyi jelentésekben ebben az esetben egy, a tanulóinak reziduális értékeit megjelenítő ábra szerepel, ami azt mutatja, hogy a telephely tanulói jobb, hasonló vagy gyengébb eredményt értek-e el, mint ami a modellben szereplő változók alapján várható lett volna. Ahogy tehát az előző hozzáadottérték-modellnél, itt sem egyetlen számmal jellemzi a jelentés a telephely eredményességét, hanem az összes tanulójának adatát közli.

Összegzés

A hozzáadottérték-modellek hasznos eszközei az oktatási intézmények eredményessége vizsgálatának, azonban alkalmazásuk megfelelő körülményeket igényel. Az elmúlt egy-két évtizedben számos tudományos cikk, tanulmány és konferencia foglalkozott mindazokkal a lehetőségekkel, felvetődő kérdésekkel, megoldandó problémákkal és módszertani részletekkel, amelyekkel a hozzáadott érték alapú értékelési rendszerek működtetőinek meg kell birkóznuk.

Tanulmányunkban áttekintettük a tanulóiteljesítmény-mérésekben alkalmazott eredménymodellek legfontosabb típusait, definiáltuk a hozzáadott érték és a hozzáadottérték-modell fogalmát. A nemzetközi szakirodalomra támaszkodva röviden bemutattuk a hozzáadott érték alkalmazásának legfontosabb szakmapolitikai és módszertani kérdéseit, melyek a hozzáadott értéket alkalmazó értékelési rendszer céljainak és következményeinek meghatározásától az adatok minőségének és mennyiségének kérdéskörén keresztül az alkalmazott statisztikai modellek megválasztásáig, illetve az eredmények megjelenítésének és értelmezésének kérdésköréig terjednek.

Ezt követően bemutattuk azokat a nagyobb hazai kutatási programokat, amelyek több időpontban is követhető módon mérik a tanulók kognitív képességeit, így adatbázisaik hozzáadott érték típusú számítások alapjául szolgálhatnak. Ilyen kutatás a Szegedi Iskolai Longitudinális Program, az Életpálya-felmérés, valamint az Országos kompetenciamérés. Utóbbi 2010 óta alkalmaz hozzáadott érték típusú elemzéseket, melyek eredményeit a telephelyi jelentésekben közli. Tanulmányunkban röviden bemutattuk, milyen statisztikai modellekkel dolgozik a kompetenciamérés, és azoknak az eredményeit hogyan építi be a jelentésekbe.

Irodalom

Balázs, I., Balkányi, P., Ostorics, L., Palincsár, I., Rábainé Szabó, A., Szepesi, I., Szipőcsné Krolopp, J., & Vadász, C. (2014). *Az Országos kompetenciamérés tartalmi keretei – Szövegértés, matematika, háttérkérdőívek*. Budapest: Oktatási Hivatal.

- Balázs, I., Lak, Á. R., Szabó, V., Szabó, L. D., & Vadász, C. (2014). *Országos kompetenciamérés 2013 – Országos jelentés*. Budapest: Oktatási Hivatal.
- Balázs, I., Ostorics, L., Szalay, B., Szepesi, I., & Vadász, C. (2013). *PISA 2012 Összefoglaló jelentés*. Budapest: Oktatási Hivatal.
- Balázs, I., & Zempléni, A. (2004). A hozottérték-index és a hozzáadott pedagógiai érték számítása a 2003-as kompetenciamérésben. *Új Pedagógiai Szemle*, 54(12), 36–50.
- Ballou, D. (2009). Test scaling and value-added measurement. *Education Finance and Policy*, 4(4), 351–383. doi: [10.1162/edfp.2009.4.4.351](https://doi.org/10.1162/edfp.2009.4.4.351)
- Ballou, D., Sanders, W., & Wright, P. (2004). Controlling for student background in value-added assessment of teachers. *Journal of Educational and Behavioral Statistics*, 29(1), 37–65. doi: [10.3102/10769986029001037](https://doi.org/10.3102/10769986029001037)
- Barber, M., & Mourshed, M. (2007). *How the world's best performing school systems come out on top*. London: McKinsey & Company.
- Betebenner, D. W., & Linn, R. L. (2010). Growth in student achievement: Issues of measurement, longitudinal data analysis, and accountability. Retrieved from <http://epa.oszk.hu/00000/00035/00036/2000-03-kf-Bognar-Value.html>
- Bognár, M. (2000). Value added analysis (a minőségbiztosítás brit eszközeinek egyike). *Új Pedagógiai Szemle*, 50(3), 99–104.
- Braun, H. (2005). Value-added modeling: What does due diligence require? In R. W. Lissitz (Ed.), *Value added models in education: Theory and application* (pp. 9–39). Maple Grove, MN: JAM Press.
- Braun, H., Chudowsky, N., & Koenig, J. A. (Ed.). (2010). *Getting value out of value-added: Report of a workshop*. Washington, DC: The National Academies Press.
- Briggs, D. C., & Betebenner, D. W. (2009, April). *Is growth in student achievement scale dependent?* Paper presented at the Annual meeting of the National Council on Measurement in Education, San Diego, CA.
- Briggs, D. C., Weeks, J. P., & Wiley, E. (2008, April). The sensitivity of value-added modeling to the creation of a vertical score scale. Paper presented at the Conference on Value-Added Modeling, Madison, WI.
- Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods. Advanced qualitative techniques in the social sciences, 1*. Thousand Oaks, CA: Sage Publications, Inc.
- Clarke, M. (2011). *Framework for building an effective student assessment system*. Washington, DC: World Bank.
- Csapó, B. (2002). Az osztályok közötti különbségek és a pedagógiai hozzáadott érték. In B. Csapó (Ed.), *Az iskolai műveltség* (pp. 269–298). Budapest: Osiris Kiadó.
- Csapó, B. (2007). Hosszmetszeti felmérések iskolai kontextusban – az első átfogó magyar iskolai longitudinális kutatási program elméleti és módszertani keretei. *Magyar Pedagógia*, 57(4), 321–355.
- Csapó, B. (2014). A szegedi iskolai longitudinális program. In J. Pál József, & Z. Vajda (Eds.), *Szegedi Egyetemi Tudástár 7. Bölcsészeti és társadalomtudományok* (pp. 117–166). Szeged: Szegedi Egyetemi Kiadó.
- Doran, H. C., & Jiang, T. (2006). The impact of linking error in longitudinal analysis: An empirical demonstration. In R. W. Lissitz (Ed.), *Longitudinal and value added models of student performance* (pp. 210–229). Maple Grove, MN: JAM Press.
- Doran, H. C., Jiang, T., Cohen, J., Gushta, M., & Phillips, G. (2005). *The precision of gain scores obtained from vertically linked scales: Implications for estimating school and teacher effects through value-added models*. Washington, DC: American Institutes for Research, Computer and Statistical Sciences Center.
- ETS, Center for K-12 Assessment & Performance Management (2010). *Exploratory seminar: Measurement challenges within the race to the top agenda*. Princeton, NJ: Educational Testing Service.
- Fitz-Gibbon, C. T. (1997). *The value added national project final report: Feasibility studies for a national system of value added indicators*. London: School Curriculum and Assessment Authority.

- Independent Evaluation Group (IEG) (2006). *From schooling access to learning outcomes: An unfinished agenda*. Washington, DC: World Bank.
- Hajdú, T., Kertesi, G., & Kézdi, G. (2014). Roma fiatalok a középiskolában. Beszámoló a TÁRKI Életpálya-felmérésének 2006 és 2012 közötti hullámaiból. In T. Kolosi, & I. G. Tóth (Eds.), *Társadalmi riport 2014* (pp. 265–302). Budapest: TÁRKI.
- Hanushek, E. A. (1971). Teacher characteristics and gains in student achievement: Estimation using micro-data. *American Economic Review*, 61(2), 280–288.
- Harris, D. N., & Sass, T. (2005, március). *Value-added models and the measurement of teacher quality*. Paper presented at the 2005 conference of the American Education Finance Association, Louisville, KY.
- Jakubowsky, M. (2008). *Implementing value-added models of school assessment*. EUI Working Papers RSCAS 2008/06. Florence: European University Institute, Robert Schuman Center for Advanced Studies.
- Kertesi, G., & Kézdi, G. (2008). A roma és nem roma fiatalok középiskolai továbbtanulása. Első eredmények a TÁRKI-Educatio Életpálya-felmérése alapján. In T. Kolosi, & I. G. Tóth (Eds.), *Társadalmi riport 2008* (pp. 344–362). Budapest: TÁRKI.
- Kertesi, G., & Kézdi, G. (2010). Iskolázatlan szülők gyermekei és roma fiatalok a középiskolában. Beszámoló az Educatio Életpálya-felmérésének 2006 és 2009 közötti hullámaiból. In T. Kolosi, & I. G. Tóth (Eds.), *Társadalmi riport 2010* (pp. 371–407). Budapest: TÁRKI.
- Kertesi, G., & Kézdi, G. (2012). A roma és nem roma tanulók teszt eredményei közti különbségekről és e különbségek okairól. *Közgazdasági Szemle*, 59(7–8), 798–853.
- Kézdi, G. (2009). *A TÁRKI Életpálya-felvétel mintavételi dokumentációja*. 4. változat. Unpublished manuscript. Retrieved from <http://econ.core.hu/file/download/eletpalya/mintavetel.pdf>
- Kolen, M. J., & Brennan, R. L. (2004). *Test equation, scaling and linking: Methods and practices*. New York: Springer. doi: 10.1007/978-1-4757-4310-4
- Ladd, H. F., & Walsh, R. P. (2002). Implementing value-added measures of school effectiveness: Getting the incentives right. *Economics of Education Review*, 21(1), 1–17. doi: 10.1016/S0272-7757(00)00039-X
- Lissitz, R. W. (Ed.). (2005). *Value added models in education: Theory and application*. Maple Grove, MN: JAM Press.
- Lissitz, R. W. (Ed.). (2006). *Longitudinal and value added models of student performance*. Maple Grove, MN: JAM Press.
- Lissitz, R. W., Doran, H., Schafer, W. D., & Willhoft, J. (2006). Growth modeling, value added modeling and linking: An introduction. In R. W. Lissitz (Ed.), *Longitudinal and value added models of student performance* (pp. 1–46). Maple Grove, MN: JAM Press.
- McCaffrey, D. F., Lockwood, J. R., Koretz, D. M., & Hamilton, L. S. (2003). *Evaluating value-added models for teacher accountability*. Santa Monica, CA: The RAND Corporation. doi: 10.1037/e658712010-001
- McCall, M. S., Kingsbury, G. G., & Olson, A. (2004). *Individual growth and school success*. Lake Oswego, OR: Northwest Evaluation Association.
- Meyer, R. H., & Dokumaci, E. (2010). Value-added models and the next generation of assessments. *ETS Center for K–12 Assessment and Performance Management*. Recuperado em, 20.
- Mourshed, M., Chijioke, C., & Barber, M. (2010). *How the world's most improved school systems keep getting better*. London: McKinsey & Company.
- OECD (2008). *Measuring improvements in learning outcomes: Best practices to assess the value-added of schools*. Paris: OECD.
- OH-KMÉO (in press). *A 2012. évi Országos kompetenciamérés technikai leírása*. Budapest: Oktatási Hivatal.
- OH-KMÉO (2014). *OKM 2013 FIT-jelentés Útmutató a Telephelyi jelentés ábráinak értelmezéséhez*. Budapest: Oktatási Hivatal.

- Patz, R. (2007). *Vertical scaling in standards-based educational assessment and accountability systems*. Washington, DC: The Council of Chief State School Officers.
- Ponisciak, P. M., & Bryk, A. S. (2005). Value-added analysis of the Chicago public schools: An application of hierarchical models. In R. W. Lissitz (Ed.), *Value added models in education: Theory and application* (pp. 40–79). Maple Grove, MN: JAM Press.
- Ray, A. (2006). *School value added measures in England*. A paper for the OECD Project on the Development of value-added models in education systems. Retrieved from: <http://webarchive.nationalarchives.gov.uk/20130401151715/http://www.education.gov.uk/publications/eOrderingDownload/RW85.pdf>
- Sanders, W. L., Wright, S. P., Rivers, J. C., & Leandro, J. G. (2009). *A response to criticisms of SAS EVAAS*. SAS Inc. Retrieved from http://www.sas.com/resources/asset/Response_to_Criticisms_of_SAS_EVAAS_11-13-09.pdf
- Sanders, W. L., Saxton, A. M., & Horn, S. P. (1997). The Tennessee value-added accountability system: A quantitative, outcome-based approach to educational assessment. In J. Millman (Ed.), *Grading teachers, grading schools: Is student achievement a valid evaluation measure?* (pp. 137–162). Thousand Oaks, CA: Corwin Press.
- Stringfield, S. C., & Yakimowski-Srebnick, M. E. (2005). Promise, progress, problems, and paradoxes of three phases of accountability: A longitudinal case study of the Baltimore city public schools. *American Educational Research Journal*, 42(1), 43–75. doi: [10.3102/00028312042001043](https://doi.org/10.3102/00028312042001043)
- Tekwe, C. D., Carter, R. L., Ma, C., Algina, J., Lucas, M., Roth, J., Ariet, M., Fisher, T., & Resnick, M. B. (2004). An empirical comparison of statistical models for value-added assessment of school performance. *Journal of Educational and Behavioral Statistics*, 29(1), 11–36. doi: [10.3102/10769986029001011](https://doi.org/10.3102/10769986029001011)
- Tóth, E. (2010). Tesztalapú elszámoltathatóság a közoktatásban. *Iskolakultúra*, (1), 60–78.
- Tóth, E., Csapó, B., & Székely, L. (2010). Az iskolák és osztályok közötti különbségek alakulása a magyar iskolarendszerben. Egy longitudinális vizsgálat eredményei. *Közgazdasági Szemle*, 57(9), 798–814.
- United Nations Educational, Scientific and Cultural Organization (UNESCO) (2007). *Education for all global monitoring report 2008: Education for all by 2015. Will we make it?* Paris: UNESCO/Oxford University Press.
- Value-added measures (2013). In S. Abbott (Ed.), *The glossary of education reform*. Retrieved from <http://edglossary.org/value-added-measures/>
- Webster, W. J. (2005). The Dallas school level accountability model: The marriage of status and value added approaches. In R. W. Lissitz (Ed.), *Value added models in education: Theory and application* (pp. 233–268). Maple Grove, MN: JAM Press.
- Webster, W. J., & Mendro, R. L. (1997). The Dallas value-added accountability system. In J. Millman (Ed.), *Grading teachers, grading schools: Is student achievement a valid evaluation measure?* (pp. 81–99). Thousand Oaks, CA: Corwin Press.
- Wright, S. P., Sanders, W. L., & Rivers, J. C. (2006). Measurement of academic growth of individual students toward variable and meaningful academic standards. In R. W. Lissitz (Ed.), *Longitudinal and value added models of student performance* (pp. 385–406). Maple Grove, MN: JAM Press.
- Wright, S. P., White, J. T., Sanders, W. L., & Rivers, J. C. (2010). SAS EVAAS statistical models. SAS Inc. Retrieved from <http://www.sas.com/resources/asset/SAS-EVAAS-Statistical-Models.pdf>
- Zumbo, B. D., & Forer, B. (2011). Testing and measurement from a multilevel view: Psychometrics and validation. In J. Bovaird, K. Geisinger, & C. Buckendahl (Eds.), *High stakes testing in education: Science and practice in K-12 settings* (pp. 177–190). Washington, DC: American Psychological Association Press. doi: [10.1037/12330-011](https://doi.org/10.1037/12330-011)

ABSTRACT

APPLYING VALUE-ADDED MODELS TO STUDENT ACHIEVEMENT

Ildikó Balácsi

The aim of this paper is to review the international literature in the rapidly developing field of value-added analysis of schools. The evaluation of the results from schools seems to be a crucial part of improving the quality of a country's school system, according to a report by McKinsey et al. (Mourshed, Chijioke & Barber, 2010). In addition, value-added analysis of schools results seems to be a better, fairer way to evaluate schools' efforts in increasing students' knowledge and competencies as compared to analysing the results of a test unconditionally, given that schools' intake varies considerably. Thus, value-added analysis of schools has become an important issue in many countries lately. However, value-added evaluation of schools is a complex and data-intensive area with many difficult policy and technical questions.

This paper lists the most common types of test-based evaluation models and then describes a commonly accepted definition of value-added models and added value of schools. It summarizes the most important issues associated with the introduction of a value-added evaluation system, from the aims and consequences of the system through the requirements on quality and quantity of data to the choice of statistical models used in the analysis.

The second part of the paper describes the largest Hungarian longitudinal student and school evaluation programmes, suitable for value-added analysis. The HELP programme in the Centre for Research on Learning and Instruction at the University of Szeged, the Hungarian Life Course Survey conducted by the Economics of Education Research Group at the Institute of Economics, and the National Assessment of Basic Competencies at the Department of Assessment and Evaluation within the Educational Authority all track students throughout their educational careers with two or more data collection events to ascertain their cognitive skills, demographics and other background variables. Based on these, the National Assessment of Basic Competencies reports three different value-added analyses for every school every year based on the reading and mathematics test booklets used in the study. The paper briefly introduces the models used to compute added value, a simple regression model at school level, a two-level (students within schools) and three-level (students within classes within schools) hierarchical linear model without background variables and a complex three-level hierarchical linear model considering various student-, class- and school-level variables in addition to students' previous results. Together with reports on students' ability scores and reports on ability scores compared to students' socio-economic background, these value-added analyses provide schools with a complex and useful description of their students' proficiencies in reading and mathematics.

Magyar Pedagógia, 116(1). 3–23. (2016)

DOI: 10.17670/MPed.2016.1.3

Levelezési cím / Address for correspondence: Balácsi Ildikó, Oktatási Hivatal, Köznevelési Programok Főosztálya, H-1055 Budapest, Szalay utca 10–14.